IN THE SPECIFICATION

Please replace the paragraph at page 28, line 12, to page 29, line 13, with the following rewritten paragraph:

FIG. 1 is a block diagram showing a video encoding apparatus which executes a video encoding method according to an embodiment of the present invention. According to this apparatus, a predictive macroblock generating unit 119 generates a predictive picture from the frame 104 stored in a first reference frame memory 117 and the frame 105 stored in a second reference frame memory 118. A predictive macroblock selecting unit 120 selects an optimal predictive macroblock from the predictive picture. A subtracter 110 generates a predictive error signal 101 by calculating the difference between an input signal 100 and a predictive signal 106. A DCT (Discrete Cosine Transform) unit 112 performs DCT for the predictive error signal 101 to send the DCT signal to a quantizer 113. The quantizer 113 quantizes the DCT signal to send the quantized signal to a variable length encoder 114. The variable length encoder 114 variable-length-encodes the quantized signal to output encoded data 102. The variable length encoder 114 encodes motion vector information and prediction mode information (to be described later) and outputs the resultant data together with the encoded data 102. The quantized signal obtained by the quantizer 113 is also sent to a dequantizer 115 to be dequantized and then to an inverse DCT unit 116. An adder 121 adds the dequantized signal and the predictive signal 106 to generate a local decoded picture 103. The local decoded picture 103 is written in the first reference frame memory 117.

Please replace the paragraph at page 29, line 20, to page 30, line 19, with the following rewritten paragraph:

In this embodiment, a local decoded picture 103 of the frame encoded immediately before the current frame is stored in the first reference frame memory 117, and a local

decoded picture 104 of the frame encoded further before the above frame is stored in the second reference frame memory 118. The predictive macroblock generating unit 119 generates a predictive macroblock signal 130, predictive macroblock signal 131, predictive macroblock signal 132, and predictive macroblock signal 133. The predictive macroblock signal 130 is a signal extracted from only the picture in the first reference frame memory 117. The predictive macroblock signal 131 is a macroblock signal extracted from only the picture in the second reference frame memory 118. The predictive macroblock signal 132 is a signal obtained by averaging the reference macroblock signals extracted from the first and second reference frame memories. The predictive macroblock signal 133 is a signal obtained by subtracting the reference macroblock signal extracted from the second reference frame memory 118 from the signal obtained by doubling the amplitude of the reference macroblock signal extracted from the first reference frame memory 117. These predictive macroblock signals are extracted from a plurality of positions in the respective frames to generate a plurality of predictive macroblock signals.

Please replace the paragraph at page 35, lines 12-25, with the following rewritten paragraph:

Referring to FIG. 5, a frame 502 is a to-be-encoded frame, and frames 500, 501, 503, and 504 are reference frames. In the case shown in FIG. 5, in encoding operation and decoding operation, the frames 500, 501, 503, 504, and 502 are rearranged in this order. In the case of encoding, a plurality of local decoded picture frames are used as reference frames. In the case of decoding, a plurality of encoded frames are used as reference frames. For a to-be-encoded macroblock 511, one of reference macroblocks 509, 510, 512, and 513 or one of the predictive signals obtained from them by linear interpolation predictions is selected on

a macroblock basis and encoded in accordance with motion vectors (505 to 508 in FIG. 5), as in the embodiment shown in FIG. 4.

Please replace the paragraph at page 38, line 20, to page 39, line 8, with the following rewritten paragraph:

As in the fifth embodiment shown in FIG. 6, a motion vector 710 with respect to the frame [[700]] 702 is encoded. A differential vector 720 between a motion vector 711 with respect to the frame 701 and the vector obtained by scaling the motion vector 710 is encoded. That is, the vector generated by scaling the motion vector 710 to 1/2 indicates a pixel 704 in the frame 701, and the differential vector 720 indicating the difference amount between the predictive pixel 705 and the pixel 704 is encoded. In general, the magnitude of the above differential vector decreases with respect to a temporally monotonous movement. Even if, therefore, the moving speed is not constant, the prediction efficiency does not decrease, and an increase in the overhead for a motion vector is suppressed. This makes it possible to perform efficient encoding.

Please replace the paragraph at page 39, line 18, to page 40, line 7, with the following rewritten paragraph:

As in the embodiment shown in FIG. 6 or 7, a motion vector 811 with respect to the reference frame 800 is encoded. A motion vector 812 with respect to the reference frame 801 can also be generated by using the motion vector obtained by scaling the motion vector 811. In the case shown in FIG. 8, however, the motion vector 811 must be scaled to 2/3 in consideration of the distance between the reference frame and the to-be-encoded frame. In the embodiment shown in FIG. 8 and other embodiments, in order to perform arbitrary scaling, division is required because the denominator becomes an arbitrary integer other than

a power of 2. Motion vectors must be scaled in both encoding operation and decoding operation. Division, in particular, requires much cost and computation time in terms of both hardware and software, resulting in increases in encoding and decoding costs.

Please replace the paragraph at page 54, line 19, to page 55, line 10, with the following rewritten paragraph:

The video encoding apparatus shown in FIG. 20 includes reference frame memories 117, 118, and 152 corresponding to the maximum reference frame count (n). Likewise, the video decoding apparatus in FIG. 21 includes reference frame memories 217, 218, and 252 corresponding to the maximum reference frame count (n). In this embodiment, in a prediction based on a linear sum, each of predictive macroblock generators 151 and 251 generates a predictive picture signal by computing the sum of the products of predictive coefficients W1 to Wn and reference macroblocks extracted from the respective reference frames and shifting the result to the right by Wd bits. The reference frames to be selected <u>by respective predictive microblock selecting units 150, 250</u> can be changed for each macroblock, and the linear predictive coefficients can be changed for each frame. A combination of linear predictive coefficients is encoded as header data for a frame, and the selection information of reference frames is encoded as header data for each macroblock.

Please replace the paragraph at page 81, lines 3-14, with the following rewritten paragraph:

In the reference frame f3 for a backward prediction for the video macroblock 61, a macroblock 60 at the same position as that of the video macroblock 61 in the frame will be considered. If a motion compensation prediction based on the linear sum of the frames F0 and F1 is used, the motion vector (<u>63 or</u> 62 in the figure) of the macroblock 60 corresponding

to the reference frame F1 for a forward prediction for the video macroblock 61 is scaled in accordance with the inter-frame distance, and the resultant vector is used as a vector for forward and backward predictions for the video macroblock 61.

Please replace the paragraph at page 82, lines 12-22, with the following rewritten paragraph:

FIG. 38 shows another example of the bi-directional prediction shown in FIG. 37. Referring to FIG. 38, a frame F0 is a reference frame for a forward prediction for a video macroblock 71 of a video frame F2, and the other arrangements are the same as those in FIG. 37. In this case, forward and backward motion vectors for the video macroblock 71 are obtained by scaling a motion vector 72 or 73 of a macroblock 70 with respect to a frame F3, which is located at the same position as that of the video macroblock 71, to the frame F0 in accordance with the inter-frame distance.

Please replace the paragraph at page 84, line 14 to page 85, line 52-22, with the following rewritten paragraph:

In this case, a motion vector with respect to one of the forward reference frames F0 and F1 for the macroblock 80 which is temporally closer to the forward reference frame F2 for the video macroblock 81 is scaled in accordance with the inter-frame distance. With this operation, forward and backward vectors for the video macroblock 81 are generated. Letting R1 be the inter-frame distance from the frame F2 to the frame F3, R2 be the inter-frame distance from the frame F4 to the frame F3, and R3 be the inter-frame distance from the frame F1 to the frame F4, a forward motion vector 84 for the video macroblock 81 is obtained by multiplying a motion vector 82 or 83 of the macroblock 80 with respect to the frame F1 by R1/R3. A backward motion vector 85 for the to-be-encoded macroblock 81 is

obtained by multiplying the motion vector 82 by -R2/R3. The video macroblock 81 is bi-directionally predicted by using the motion vectors 84 and 85 obtained by scaling.

Please replace the paragraph at page 86, lines 4-15, with the following rewritten paragraph:

Of to-be-encoded pixel blocks A, B, C, and D in the video frame F3, for the blocks A, B, and C, reference pixel block signals with motion compensation are generated from <u>reference blocks 90, 91, 92 in</u> the frames F1, F0, and F2, respectively. With respect to these reference pixel block signals, a prediction pixel block signal is generated by multiplications of weight factors and addition of DC offset values. The difference between the prediction pixel block signal and the to-be-encoded pixel block signal is calculated, and the differential signal is encoded, together with the identification information of the reference frames and motion vector information.